

WHAT IS CLAIMED IS:

- 1 1. A method of consolidating genealogy records, comprising:
2 partitioning the records using at least one index file to form one or more
3 partitions;
4 sorting the records in a partition based on a data element in the records;
5 comparing records within a sort range;
6 based on the comparison, identifying same person records;
7 consolidating information in the same person records;
8 receiving a request from a user to view at least a portion of the consolidated
9 information for a particular group of same person records; and
10 sending a file comprising the portion to the user.
- 1 2. The method of claim 1, wherein partitioning the records using at least
2 one index file comprises using a surname index to identify records having the same surnames
3 and grouping those records into a surname partition.
- 1 3. The method of claim 2, further comprising using the surname index to
2 identify records having similar surnames and grouping those records into the surname
3 partition.
- 1 4. The method of claim 3, wherein using the surname index to identify
2 records having similar surnames comprises using a phonetic algorithm to identify records
3 having similar surnames.
- 1 5. The method of claim 4, wherein the phonetic algorithm comprises
2 double metaphone.
- 1 6. The method of claim 4, wherein the phonetic algorithm comprises
2 SOUNDEX.
- 1 7. The method of claim 1, wherein sorting the records in a partition based
2 on a data element in the records comprises sorting the records based on birth date.
- 1 8. The method of claim 1, wherein sorting the records in a partition based
2 on a data element in the records comprises sorting the records based on a selection from the
3 group consisting of name, death data, death place, and birth place.

1 9. The method of claim 7, wherein comparing records within a sort range
2 comprises comparing records within a birth date range.

1 10. The method of claim 1, wherein identifying same person records
2 comprises calculating a score that represents the likelihood that a pair of compared records
3 represent the same person.

1 11. The method of claim 10, further comprising comparing records related
2 to pairs of same person records.

1 12. The method of claim 11, wherein comparing records related to pairs of
2 same person records comprises revising the score based on the comparison of related records.

1 13. The method of claim 12, wherein identifying same person records
2 comprises comparing the score to a predetermined threshold and rejecting records as “same
3 person” records if the score is below the threshold.

1 14. The method of claim 1, wherein the portion comprises a family tree
2 based on consolidated information from a plurality of records.

1 15. A system for consolidating genealogy records, comprising:
2 a processor programmed to:
3 partition the records using at least one index file to form one or more
4 partitions;
5 sort the records in a partition based on a data element in the records;
6 compare records within a sort range;
7 based on the comparison, identify same person records;
8 consolidate information in the same person records;
9 receive a request from a user to view at least a portion of the
10 consolidated information for a particular group of same person records; and
11 send a file comprising the portion to the user.

1 16. The system of claim 15, wherein the processor, in being programmed
2 to partition the records using at least one index file is further programmed to use a surname
3 index to identify records having the same surnames and grouping those records into a
4 surname partition.

1 17. The system of claim 16, the processor is further programmed to use the
2 surname index to identify records having similar surnames and group those records into the
3 surname partition.

1 18. The system of claim 17, wherein the processor, in being programmed
2 to use the surname index to identify records having similar surnames is further programmed
3 to use a phonetic algorithm to identify records having similar surnames.

1 19. The system of claim 18, wherein the phonetic algorithm comprises
2 double metaphone.

1 20. The system of claim 18, wherein the phonetic algorithm comprises
2 SOUNDEX.

1 21. The system of claim 15, wherein the processor, in being programmed
2 to sort the records in a partition based on a data element in the records is further programmed
3 to sort the records based on birth date.

1 22. The system of claim 15, wherein the processor, in being programmed
2 to sort the records in a partition based on a data element in the records is further programmed
3 to sort the records based on a selection from the group consisting of name, death data, death
4 place, and birth place.

1 23. The system of claim 21, wherein the processor, in being programmed
2 to compare records within a sort window is further programmed to compare records within a
3 birth date range.

1 24. The system of claim 15, wherein the processor, in being programmed
2 to identify same person records is further programmed to calculate a score that represents the
3 likelihood that a pair of compared records represent the same person.

1 25. The system of claim 24, further the processor is further programmed to
2 compare records related to pairs of same person records.

1 26. The system of claim 25, wherein the processor, in being programmed
2 to compare records related to pairs of same person records is further programmed to revise
3 the score based on the comparison of related records.

1 27. The system of claim 26, wherein the processor, in being programmed
2 to identify same person records is further programmed to compare the score to a
3 predetermined threshold and reject records as “same person” records if the score is below the
4 threshold.

1 28. The system of claim 15, wherein the portion comprises a family tree
2 based on consolidated information from a plurality of records.